SCOPE OF WORK

TITLE: Metrics and Indicators of Final Ecosystem Goods and Services: Identification of Ecosystem Goods and Services in Support of Benefits Analysis

1. Background

The purpose of the technical support requested in this work assignment, consistent with sections G3 and I23 of the contract's Statement of Work, is to make quantum improvements in the nation's capacity to identify and utilize biophysical measures and indicators that are both responsive to air pollution (and other stressors) and which contribute most clearly to human well-being.

A. Air Pollution and Ecosystems

The detrimental effects of air pollution on ecological resources has been the subject of research for centuries (Cowling 1982). Focused efforts over the last few decades have developed and used the capacity to construct national and international assessments of air pollutants. These syntheses identify ecological impacts ranging from degradation of soils, damage to forests and crops, shifts in plant community composition, acidification and eutrophication of surface waters with consequent effects on their biota, and increases of nitrate in ground water to levels that exceed drinking water standards. Decisions about the management of this stress reflect multiple considerations including the benefits associated with reductions in emissions. A recent analysis of the benefits of a major national effort to manage air pollution included an analysis of ecological benefits, but noted:

> "…quantitative assessment remains problematic due to a lack of units of measure to gauge changes in the quality and quantity of ecosystem services…" (Chestnut and Mills 2005).

B. Ecosystems and Human Well Being

Management of natural resources benefits from data to support a wide range of basic and practical public purposes. For example,

1. They can provide fundamental insights into the ways in which ecosystems function; this understanding can lead to the construction of predictive models.
2. They can document the extent to which we comply with or make progress towards compliance with regulations.
3. They can tell us if people getting what they want from ecosystems and thus provide the foundation for the analysis of social well-being.

While purposes such as these are not independent of one another, there are distinct practices and sets of skills required for defining the data suitable for each. Natural scientists have focused their efforts on questions such as the first two, and in acknowledging the third purpose they have also recognized the need to design the data specifications for this goal in partnership with social scientists. In fact, provoked by the work of social scientists such as Boyd and Banzhaf (2007), teams of natural and social scientists have begun to translate social science principles into practices that lead to the identification of the kinds of natural resources data most useful for the analysis of social well being. These teams have been organized within EPA's Ecosystem Service Research Program (ESRP). It is the purpose of this agreement to extend that work and the collaboration that supports it.

The ESRP Monitoring Program, through the Freshwater Ecology Branch (FEB) in EPA ORD's Western Ecology Division is designing a national program reporting on indicators of final ecosystem goods and services (FEGS). FEGS are defined as biophysical features, quantities and qualities requiring little further translation to make clear their relevance to human well being (Boyd and Banzhaf 2007). They are the components of ecosystems perceived by people to be directly relevant to their welfare, as opposed to the larger set of ecological components on which the final goods and services depend. These are the ecological units that serve as the foundation for the analysis of social well-being. It is necessary to specify these units not only for national monitoring programs, but also for monitoring at other scales, and for the development of ecological production function models necessary to support analysis of social well being (e.g. Chee 2004; Daily and Matson 2008).

FEB has begun to extend the FEGS concepts into a practical set of metrics. This extension has been developed in two workshops attended by natural and social scientists. The two key results of these workshops has been 1) the identification of a candidate set of biophysical metrics of FEGS for three aquatic ecosystems -- streams, wetlands and estuaries – and 2) the development of a process to enable the identification of candidate metrics of FEGS for other systems.

C. Key Questions

This background leads to four key questions.

1) Has the identification of metrics of FEGS for these three ecosystems been sufficient? (Metrics)
2) What refinements, if any, should be made to the process developed for identifying metrics of FEGS as it may be applied to other ecosystems (Process)
3) How do the metrics aggregate to indicators of FEGS at national scales and what are the implications of that aggregation for the specification of metrics? (Aggregation)
4) Do national and regional static or dynamic models predict changes in metrics in response to changes in atmospheric loads and levels of Nr and SOx? (Air)

These questions structure the tasks the EPA wishes the contractor to undertake.

Question 1: Has the identification of metrics for three ecosystems been sufficient?

EPA's efforts (Ringold, Boyd et al. 2009; Ringold and Landers 2010; Ringold, Boyd et al. In Review) have harnessed the expertise of dozens of natural and social scientists to identify candidate metrics of FEGS for specific ecosystems. These efforts identified candidate metrics with three level of specificity. In some instances these workshops provided the specificity that would allow one to rigorously match requirements to a measurement protocol or a model prediction. In other instances the workshop provided a qualitative statement (e.g. water with pathogen levels safe for swimming) without providing much specificity. In other cases, especially for aesthetics, we were only able to identify that research (or expertise beyond that of the workshop participants) would need to be conducted to identify candidate metrics. Even in cases where workshop outputs provided the most specificity we recognize that workshop results provide a reasonable working hypothesis that would benefit from empirical evaluation.

In parallel to the biophysical specification of FEGS metrics, there is a need to define the sample unit for these metrics and the temporal and spatial dimensions of the sample unit. While there are well established procedures for determining the dimensions of a biophysical sample unit sufficient for ecological analysis, we have not been able to identify a procedure for determining the dimensions of a biophysical sample unit for analyses of social well being. How big is the biophysical unit valued by a catch and release angler? What is the spatial unit that should be sampled that would provide meaningful information for a subsistence hunter? What is the sensible temporal unit of sampling to represent a resource for non-use benefits? While we recognize that such dimensions may not frequently exist for use in social analysis, we seek guidance on how to proceed in the absence of such specification.

In response to our first question we are interested in the review of our existing work resulting in revised checkmark matrices and metric matrices in a form similar to that provided in (Ringold, Boyd et al. 2009; Ringold, Boyd et al. In Review). Our expectation is that considerable deference shall be given to the judgments made at the workshops unless empirical evidence identifies a markedly different result. Any recommendations for revisions to these two matrices, along with the rationale for the revision shall be provided as part of Deliverable 2. In addition, we seek review about the sufficiency of specification of each metric. Is it described with biophysical specificity sufficient to be implemented -- largely a natural science task. In parallel, the effort shall define and illustrate how one determines the sufficiency of the specification of the temporal and spatial dimensions of the metric. This effort will address questions such as what is the temporal and spatial unit that creates value for beneficiaries. This work may focus on major groupings or categories of beneficiaries rather than all beneficiaries. In addition, when the answers to questions such as these are not clearly known, the analysis will recommend how monitoring and modeling should proceed in the presence of this uncertainty.

For each metric sufficiently specified the analysis shall describe the feasibility and likelihood of a substantial impact on human well being. Metrics that are unlikely to have a substantial effect on human well being, whether positive or negative shall be identified. This analysis shall consider not only the likely effect of a metric on human well being, but also the likely cost-effectiveness of providing information on the metric. This analysis shall be based on human well being in the aggregate, rather than with regard to a specific single beneficiary or group of beneficiaries. The analysis shall classify the remaining metrics into four classes of feasibility for implementation in a national or regional monitoring or modeling effort: Currently feasible, feasible in the short-term, feasible in the long-term and unlikely to be feasible. This classification shall consider cost-effectiveness in their classification of which metrics are feasible. Based on the answer to this question the effort shall undertake two additional efforts. The first is to conduct a

gap analysis of existing large scale monitoring and modeling capacity in light of the list of metrics considered likely to have a substantial effect on human well being. At a minimum this GAP analysis shall focus on the NARS programs for streams, wetlands and estuaries and on the TIME and LTM stream monitoring programs. This analysis will be useful because it will enable us to compare and contrast programs with two different temporal and spatial characteristics. The NARS programs (http://water.epa.gov/type/watersheds/monitoring/nationalsurveys.cfm) have a national spatially extensive design focusing broadly on the biotic integrity goals of the Clean Water Act; the TIME/LTM program is a regional program focusing on regional responses to changes in acidic deposition with a temporal record in excess of twenty years. This contrast is important because a preliminary gap analysis identified issues of temporal and spatial scales as one of the most significant barriers to a national ecosystem services monitoring program (Ringold, Boyd et al. In Review). The goal of this gap analysis shall be to help identify the magnitude of the gap in terms of 2 factors: 1) the likely consequence of the gap for providing analyses of human well being, 2) the obstacles to adding the metric to large scale monitoring and modeling programs. This analysis shall illustrate the gap in practical terms with existing data or existing model results. Finally, the analysis shall make recommendations describing the highest priority practical measures for change in national monitoring and modeling programs. In addition, the effort shall recommend highest priority research areas.

Question 2: What refinements, if any, should be made to the process we have developed as it may be applied to other ecosystems?

Our efforts to identify a set of metrics for three aquatic ecosystems, was based on a process developed and refined during the workshops. As we seek a comprehensive set of metrics of FEGS for all ecosystems we would transfer the process used for these three ecosystems to other ecosystems. Our process, described in more detail in the works noted above, consists of four steps all based on the judgment of groups of experts rooted in multiple disciplines:

1. Define ecosystem boundaries
2. Identify beneficiaries of the ecosystem's goods and services and the broad attributes of the ecosystem that provide those goods and services.
3. Identify the attributes providing a final good or service for each beneficiary (See the column headings of Table 1 on page 22 in (Ringold, Boyd et al. 2009)).
4. Identify metrics for each attribute providing a final good or service.

Following these steps we've made considerable progress. We seek focused input from additional experts on whether and how to improve this process.

Question 3: How do the metrics aggregate to indicators of FEGS at national scales and what are the implications of that aggregation for the specification of metrics?

We have identified metrics for individual beneficiaries. However, policy, and the assessment of human well being upon which policy wisely relies, should be formulated for multiple individuals over large areas and long periods. Thus the metrics, to be useful, require several different types of aggregation as illustrated in Figure 1. The key question for us is to identify what implications approaches to aggregation may have for metric identification or priorities.

Multiple metrics of FEGS were identified for each beneficiary. Combinations of these metrics provide an indicator of a final good or service. For example, water quantity and water chemistry, especially conductivity, combine to create an indicator of the FEGS provided for an irrigator. If we know how those two metrics combine at a point in time for a specific beneficiary how do we aggregate that information over large areas and long periods of time. How does that reporting differ for rival goods (goods, such as fish or water, whose consumption by one user prevents consumption by another user) and compared to non-rival goods. Most importantly, in this aggregation process what are the implications for the way metrics are specified? Is there any opportunity for simplification in that specification?

The effort shall consider how to set priorities for approaching the aggregation issues illustrated in Figure 1. In the development of these priorities and analysis of these issues the effort shall consider the views and ideas developed by Ringold and Landers (2010). In addition, as examples are illustrated, and gaps are identified, the effort shall identify and illustrate the implications of the aggregation for the selection and specification of metrics.

The first aggregation issue to be addressed is how to aggregate multiple FEGS metrics into an indicator of a FEGS for a beneficiary. The effort shall a) identify and demonstrate existing approaches linking multiple metrics with indicators of human well being for individual beneficiaries, and b) when existing approaches are inadequate, the effort shall identify priorities for specific research to link multiple metrics to indicators of human well being.

Second, the effort shall identify and illustrate existing approaches in which biophysical metrics and indicators of FEGS can be effectively aggregated from individual beneficiaries to aggregations of beneficiaries embodied in individuals or organizations or effective groupings of individuals and organizations. In identifying the efficacy of existing approaches, the effort shall also report on gaps in the capacity of existing approaches and shall identifying approaches to resolving the most important gaps. This effort shall explicitly consider the “community approach” described by Ringold and

Landers(2010). The community approach focuses on aggregations of behaviorally similar people and the ecological features they value as a means to aggregate ecosystem values to populations of people as a whole rather than by aggregating from individual beneficiaries. T

Third, the effort shall identify and illustrate existing approaches in which biophysical metrics and indicators of FEGS can be effectively aggregated from sample units to larger areas (i.e. assessment units). In identifying the efficacy of existing approaches, the effort shall also report on gaps in the capacity of existing approaches and shall identify approaches to resolving the most important gaps.

Question 4: Do national and regional static or dynamic models predict or describe changes in FEGS metrics in response to changes in atmospheric loads and levels of Nr and SOx? (Air)

Analysis of human well being depends on the analysis of how incremental change in a stressor leads to incremental change in FEGS. In practice this requires models (in ecosystem services taxonomy these are referred to as production functions because they are analogous to and linked to economic production functions -- (Boyd and Krupnick 2009)) to enable predictions of this relationship. Evaluation of the capacity of models to provide this information for air pollution and the "substantial" metrics for aquatic ecosystems identified in these analyses can benefit the design of future models and illustrate this process for other systems. Given the sustained attention to constructing regional, national and international data and modeling systems to address this issue, the air pollution ecosystem system is an excellent prototype. To support this analysis the following questions must be addressed:

First, which of the "substantial" metrics (Question 7 in Table 2) plausibly respond to atmospheric exposure. To make this analysis meaningful it should focus on the range of exposures currently observed in the United States. Second, which of these "sensitive" and "substantial" metrics are reasonably estimated in currently operational static and dynamic regional or national models linking ecosystems to changes in deposition of Nr or SOx?

Task 1 Evaluation of Existing Work and Input to Design of Future Work

The contractor shall identify two social scientists to participate with natural scientists from EPA as a core group to fully engage in tasks 1 and 2 of this work assignment. One member of the core group shall have

firsthand knowledge of the planning and implementation of (Ringold et al. 2009, Ringold et al. In Review); the second member shall have conceptual knowledge and practical experience in designing, implementing and analyzing human preferences using diverse quantitative and qualitative methods. The core group shall work with EPA scientists in tasking the expert reviewers and in making recommendations to EPA about the activities to be conducted under Task 2. The contractor shall identify up to seven scientists to participate in an interdisciplinary review of the materials developed by this ESRP effort to date (to be provided by the government) and the questions listed above. The specific charge to the reviewers will be prepared by the full core group as Deliverable 3.The experts shall be identified by the contractor to reflect the breadth of expertise required to address these questions. This breadth includes 1) economists intimately familiar with the final ecosystem services concept as represented in (Boyd and Banzhaf 2007) and applied in (Ringold, Boyd et al. 2009; Ringold, Boyd et al. In Review; Ringold, Boyd et al. In Review), 2) social scientists conversant with how to efficaciously measure, sample and aggregate human values and perspectives, and 3) natural scientists with operational familiarity with the principles of design for monitoring ecosystems at national scales. The list of experts shall be submitted to the government for review to ensure their individual and collective capacity to respond to the technical needs embodied in the questions. This list is Deliverable 4 and 5. After receiving EPA approval, the contractor shall secure reviews by the experts of the work completed and suggest approaches for pursuing the questions listed below and summarized in Table 2. The reviews shall take two forms:

1) In direct form the experts shall address the work done, e.g. in the view of the expert is each metric reasonably specified? Or
2) In procedural form the experts shall identify approaches and people to address the questions directly. For example, how to we go about determining the cost-effectiveness of providing information about each metric?

These written reviews and recommendations shall constitute Deliverable 6. RTI shall then prepare a draft synthesis of the reports (Deliverable 7) and circulate it among the experts for their review. This report shall be finalized and provided as Deliverable 8. Deliverable 8 shall be circulated among the other core group (including EPA personnel in advance of a telephone or, if possible a video based “workshop”. The purpose of the workshop will be to the core group to discuss refinements to EPA questions posed below and efficient approaches for addressing them.

The contractor shall provide support for the core group to participate in the "workshop" and for all logistical support at the workshop. The contractor shall provide personnel at the workshop to manage the logistics, to facilitate the discussion, and to maintain a record of the highlights of the discussion. This record of highlights and recommendations constitutes Deliverable 9. In addition, workshop participants shall reach consensus on the entries in the matrices from previous workshops. Their conclusions shall represent Deliverable 9. EPA will consider these views and then, potentially, will issue a work assignment amendment to pursue those recommendations or other recommendations that in EPA's view best enable EPA to address the key questions listed above. It is expected that the core group providing Deliverable 9 shall stay involved in the next set of activities.

Task 2 Additional Workshops (Optional)

The contractor shall provide support for additional workshops. The provision of Deliverable 9 will lead EPA to issue a request for additional deliverables to address the four key questions (See page 3) unresolved but adequately focused by the initial review of the experts. For planning purposes it is expected that progress will take the form of two workshops the first attended by twenty scientists at a cost-effective facility to be proposed by RTI; the second attended by ten scientists at a similar facility. Note that while the form or process by which these questions shall be addressed may not be a workshop, it is the intention of the government to direct the contractor to pursue these questions in an efficacious manner. The contractor shall include these workshops in the work plan and cost estimate, however, no effort shall be expended on this task until explicitly directed by an amendment to this work assignment.

Task 3 Goods vs Services

The contractor shall designate an economist intimately familiar with the final ecosystem services concept as represented in (Boyd and Banzhaf 2007) and applied in (Ringold, Boyd et al. 2009; Ringold, Boyd et al. In Review; Ringold, Boyd et al. In Review) to provide technical support for an evaluation of the distinction between goods and services. The Ecosystem Services literature uses "Ecosystem Services" as a short hand term for Ecosystem Goods and Services without recognition of the distinction well developed in the social science literature between goods and services (e.g. Lovelock and

Gummesson 2004; Kotler and Keller 2009). EPA intends to evaluate this distinction and how it relates to the further development of its concept of FEGS. This designation shall take the form of Deliverable 12A. Having secured EPA concurrence, the form of this technical support shall be contributions to and comments on a manuscript to be submitted to a peer reviewed journal prepared by an EPA scientist. These contributions shall be summarized in Deliverable 12B. The contractor shall include these workshops in the work plan and cost estimate, however, no effort shall be expended on this task until explicitly directed by an amendment to this work assignment.

PAGE INTENTIONALLY LEFT BLANK

Table 1. List of requested Deliverables

Task	Deliverable	Description	Suggested Due Date
1	1	Proposed List of Core Group Members	One month after approval of workplan
1	2	Final List of Core Group Members	Two weeks after WACOR review of Deliverable 1
1	3	Charge to reviewers	Two weeks after provision of draft from WACOR
1	4	Proposed List of up to Seven Experts	Two weeks after Deliverable 2
1	5	Final List of Up to Seven Experts	Two weeks after comments from WACOR and Core Group members on Deliverable 4
1	6	Individual Expert Analyses	Six weeks after approval of the list of experts
1	7	Draft Synthesis report prepared by RTI including necessary and documented revisions to the checkmark and metric matrices	One month after deliverable 6
1	8	Final synthesis report including necessary and documented revisions to the checkmark and metric matrices	Six weeks after Deliverable 7
1	9	Record of discussion of Core Group Members on Deliverable 8	Within one month after Deliverable 8.
2	10	Report from the second workshop	To be specified in an amendment to the work assignment
2	11	Report from the third workshop	To be specified in an amendment to the work assignment
3	12 A and B	A. Proposed technical expert B. Technical contributions to an evaluation of the distinction between goods and services and its relevance to the development and application of FEGS	A. One month after approval of the workplan B. Status reports every other month once proposed expert is approved.

Table 2. Questions to be addressed in this work assignment. See Text

Question Class	Specific Question	Approach to Address in Deliverable 2
Metrics	1. Is each metric reasonably specified?	Directly and Procedurally
Metrics	2. Is each metric specified sufficiently for implementation?	Directly and Procedurally
Metrics	3. Are the temporal and spatial dimensions of the metrics reasonably well known for use in analyses of well being?	Directly and Procedurally
Metrics	4. How should monitoring and modeling proceed if answers about temporal and spatial dimensions are ambiguous?	Directly and Procedurally
Metrics	5. What is the likelihood that each metric has a substantial effect on human well being?	Directly and Procedurally
Metrics	6. What is the probable cost-effectiveness of providing information on each "substantial" metric?	Procedurally
Metrics	7. What is the feasibility of including each "substantial" metric in a national monitoring program?	Procedurally
Metrics	8. What is the "gap" between substantial metrics and current national (including NARS) and regional (including TIME/LTM) monitoring programs?	Procedurally
Metrics	9. What are the highest priorities metrics for inclusion in national and regional monitoring programs	Procedurally
Process	10. How can the process for identifying FEGS be improved	Directly
Aggregation	11. How do FEGS metrics aggregate meaningfully for a single beneficiary?	Procedurally
Aggregation	12. Are there approaches to identify coherent groups of people for which a parsimonious set of metrics sensibly link to human well being?	Procedurally
Aggregation	13. How do types of FEGS indicators aggregate over time and space?	Procedurally
Air	14. Is each substantial metric plausibly affected by atmospheric deposition of Nr or Sox or exposure to other air pollutants at loads or levels currently found in the United States?	Procedurally
Air	15. Which substantial and sensitive metrics are reasonably estimated in currently operational static and dynamic regional or national models linking ecosystems to changes in atmospheric deposition of or exposure to Nr or SOx?	Procedurally

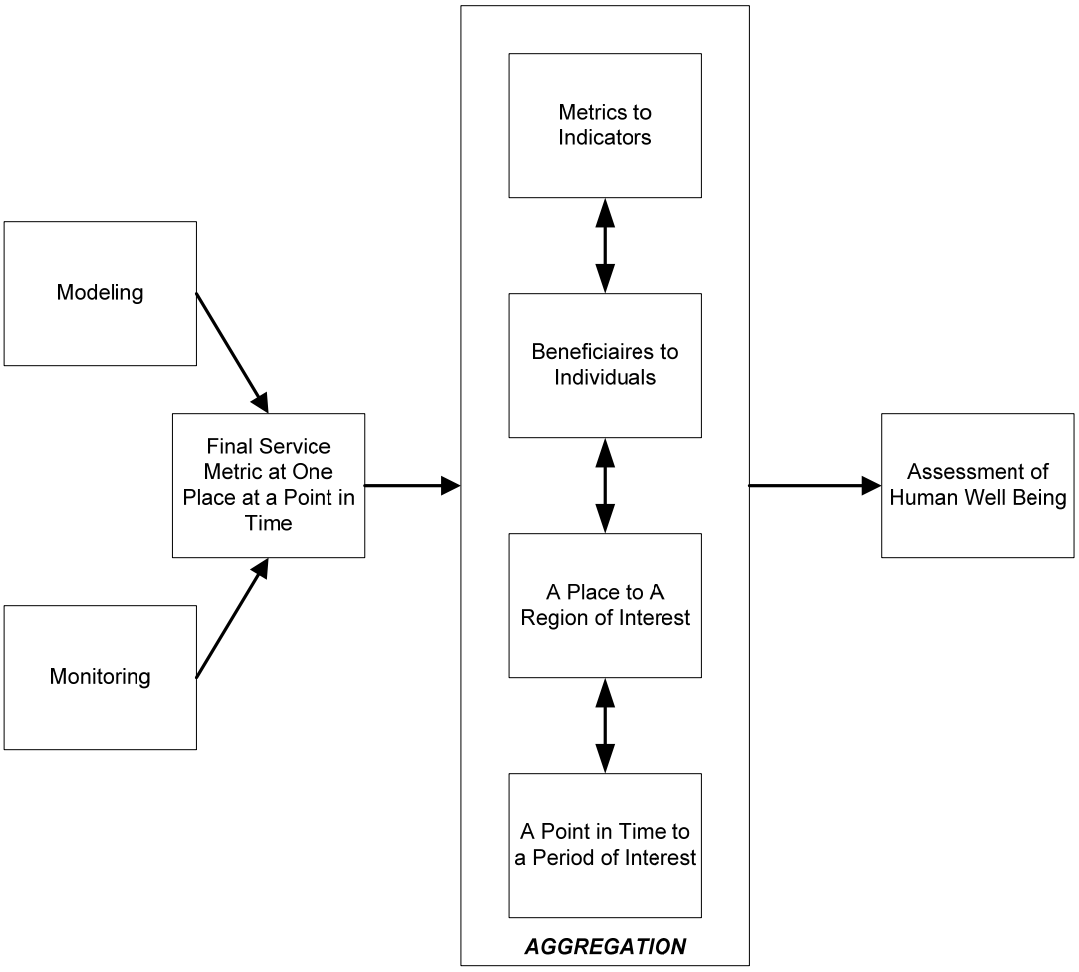

Figure 1. Illustration of the types of aggregation required to convert metrics into assessments of human well being.

List of Literature Cited

Boyd, J. and S. Banzhaf (2007). "What are ecosystem services? The need for standardized environmental accounting units." Ecological Economics **63**(2-3): 616-626.

Boyd, J. W. and A. J. Krupnick (2009). The Definition and Choice of Environmental Commodities for Nonmarket Valuation. Washington, D.C., Resources for the Future**:** 60.

Chee, Y. E. (2004). "An ecological perspective on the valuation of ecosystem services." Biological Conservation **120**(4): 549-565.

Chestnut, L. G. and D. M. Mills (2005). "A fresh look at the benefits and costs of the US acid rain program." Journal of Environmental Management **77**(3): 252.

Cowling, E. B. (1982). "Acid precipitation in historical perspective." Environmental Science & Technology **16**(2): 110A-123A.

Daily, G. C. and P. A. Matson (2008). "Ecosystem services: From theory to implementation." Proceedings of the National Academy of Sciences **105**(28): 9455-9456.

Kotler, P. and K. L. Keller (2009). Designing and Managing Services. Marketing Management. Upper Saddle River, NJ, Pearson Prentice Hall.

Lovelock, C. and E. Gummesson (2004). "Whither Services Marketing?" Journal of Service Research **7**(1): 20-41.

Ringold, P. L., J. W. Boyd, et al. (2009). Report from the Workshop on Indicators of Final Ecosystem Services for Streams. U. S. E. P. Agency. Corvallis, OR**:** 56.

Ringold, P. L., J. W. Boyd, et al. (In Review). "A Framework for Identifying Indicators of Ecosystems Contributions to Human Well Being: A Case Study with Streams."

Ringold, P. L., J. W. Boyd, et al. (In Review). Report from the Workshop on Indicators of Final Ecosystem Services for Wetlands and Estuaries. U. S. E. P. Agency. Corvallis, OR**:** 56.

Ringold, P. L. and D. H. Landers (2010). A Draft Monitoring Strategy for the USEPA Ecosystem Services Research Program. O. o. R. a. D. US EPA, Western Ecology Division. Corvallis, OR**:** 29.